A Mathematical Formalism for the Topic Maps Reference Model

Neill A. Kipp

DRAFT: October 14, 2003

1 Purpose

A **Topic Map** is a deterministic formal system of elements and sets of elements that models conversation and records knowledge. The set of axioms herewith explicitly defines topic maps, submaps, and merged maps. The satisfaction of all the axioms by a candidate set means that the set is a topic map.

2 Set Theory

Set theory provides the basis for our formal notation. Readers familiar with set theory can safely skip to the next section.

A **set** is a collection of distinguishable objects called **elements**. If x is an element of set A, we write $x \in A$. If x is not an element of set A, we write $x \notin A$. A set cannot contain the same element more than once. We can describe a set by listing its elements in braces, $\{1, 2, 3\}$, or by describing its contents, {all houses on my street}. The elements of a set are not ordered. Two sets are **equal** if they contain the same elements, for example $\{1, 2, 3\} = \{3, 1, 2\}$. Sets are not equal otherwise.

Sets can contain other sets, as in {{house, apartment}, {street, avenue}}, but sets cannot contain themselves. Therefore the "set of all sets" is not allowed. Also, elements of a set need not have the same type or be similar in any way; thus {bicycle, Springfield, $19.95} is a set.

If all the elements of set A are also elements of set B, then we say A is a **subset** of B, and write $A \subseteq B$. The **empty set** $\{\}$ is a set that has no elements. The empty set is a subset of all sets, including itself. The set of all subsets of A is the **power set** of A, written 2^A. For example,

$$2^{\{1,2,3\}} = \{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

The number of elements in a set is written $|A|$. Therefore $|\{\}| = 0$, $|\{0\}| = 1$, $|\{0, 1\}| = 2$, and $|2^A| = 2^{|A|}$.

The **intersection** of two sets is written $A \cap B$, and is defined as $\{x : x \in A \text{ and } x \in B\}$, which is read, "the intersection of A and B is the set of all elements x such that x is an element of A and x is an element of B." The **union** of two sets is $A \cup B = \{x : x \in A \text{ or } x \in B\}$.

An **ordered pair** of two elements a and b is written (a, b). An ordered pair is formally defined as $(a, b) = \{a, \{a, b\}\}$. Thus the ordered pair $(1, 2)$ is not the same as the ordered pair $(2, 1)$. The **Cartesian product** of two sets A and B is $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$. For example,

$$\{0, 1\} \times \{0, 1, 2\} = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2)\}.$$

Similarly the Cartesian product of three sets $A \times B \times C$ is a set of ordered triples, and the Cartesian product of n sets $A_1 \times A_2 \times A_3 \times ... \times A_n$ is a set of ordered n-tuples.

Any subset of $A \times B$ is called a **binary relation** on A and B. For example, let $A = \{0, 1\}$, let $B = \{(0, 1, 2)\}$ and let $R = \{(0, 0), (0, 1), (1, 2))\}$. Because $R \subseteq A \times B$, R is a binary relation on A and B. Indeed an n-place relation is a subset of a Cartesian product with n operands; n-place relations contain n-tuples.

A **function** f is a binary relation on A and B such that for all $a \in A$, there exists at most one $b \in B$ where $(a, b) \in f$. Notation $f : A \to B$ is read, "set f is a function from A to B." Herein we write $\{f : A \to B\}^*$ to mean the set of all possible functions f from A to B. As a shorthand we write $f(a) = b$ instead of $(a, b) \in f$.

The following symbols have these meanings:

Symbol	**Meaning**
$\wedge$	and
$\vee$	or
$\Rightarrow$	implies
$\Leftrightarrow$	if and only if
$\exists$	there exists
$\nexists$	there does not exist
$\forall$	for all

3 Axioms

A Topic Map (M) is an ordered six-tuple whose elements denote the types (T), roles (R), assertions (A), castings (C), other related elements (Z), and independent elements (I) of conversation.

$$M = (T, R, A, C, Z, I). \tag{1}$$

The elements of M are disjoint. Types cannot be roles, assertions, castings, other related elements, or independent elements. Roles cannot be assertions, castings, other related elements, or independent elements. Assertions cannot be castings, other related elements, or independent elements. Castings cannot be other related elements or independent elements. Other related elements cannot be independent elements.

$$\begin{aligned} T \cap R = T \cap A = T \cap C = T \cap Z = T \cap I &= \\ R \cap A = R \cap C = R \cap Z = R \cap I &= \\ A \cap C = A \cap Z = A \cap I &= \\ C \cap Z = C \cap I &= \\ Z \cap I &= \{\}. \end{aligned} \tag{2}$$

The set of types is a subset of all the possible subsets of the set of roles. In other words, elements of the set of types are themselves sets, where each type is a set of roles.

$$T \subseteq 2^R. \tag{3}$$

Each type must have two or more roles.

$$t \in T \Rightarrow |t| \geq 2. \tag{4}$$

A role cannot appear in more than one type. In other words, the intersection of any two distinct types is the empty set.

$$(t_i \in T \wedge t_j \in T \wedge t_i \neq t_j) \Rightarrow (t_i \cap t_j = \{\}). \tag{5}$$

The set of roles is the union of all the elements of all the types.

$$R = \bigcup_{\forall t_i \in T} t_i. \tag{6}$$

The set X is the set of dependent elements.

$$X = T \cup R \cup A \cup C \cup Z. \tag{7}$$

Each assertion has a type, roles, and role players. Each assertion is an ordered pair whose first member is a type t and whose second member is a function f that denotes a mapping from roles to role players in the assertion. Independent elements (I) cannot be role players in assertions.

$$A \subseteq T \times \{f : R \to X\}^*. \tag{8}$$

An assertion must have at least one role player. In other words, for each $(t, f) \in A$, the function f must have at least one element.

$$(t, f) \in A \;\Rightarrow\; |f| \geq 1. \tag{9}$$

If a role appears in an assertion, it must appear in the assertion's type.

$$(t, f) \in A \wedge f(r) = x \;\Rightarrow\; r \in t. \tag{10}$$

Castings relate assertions, roles, and possibly role players. A casting can be an ordered triple or an ordered pair.

$$C \subseteq (A \times R \times X) \cup (A \times R). \tag{11}$$

Castings relate an assertion, role, and role player when the assertion has a role player for the role. Castings relate an assertion to a role when the assertion does not have a role player for the role. No castings exist otherwise.

$$\begin{aligned} C \;&=\; \{(a, r, x) : \forall a \in A, a = (t, f) \wedge f(r) = x\} \\ &\quad\cup \{(a, r) : \forall a \in A, a = (t, f) \wedge \nexists x \text{ such that } f(r) = x\}. \end{aligned} \tag{12}$$

The set of other related elements are the role players that are not also types, roles, assertions, or castings.

$$Z = \{x : \forall (a, r, x) \in C \wedge x \notin (T \cup R \cup A \cup C)\}. \tag{13}$$

4 Useful Topic Map Operations

In the following, we define set M^* to be the set of all topic maps. In other words, if J is a topic map, then J is an element of M^*.

The submap relation is a binary relation on topic maps.

$$\text{submap} \subseteq M \times M. \tag{14}$$

Topic map J is a **submap** of topic map K if and only if the all elements of J are subsets of the corresponding elements of K.

$$\begin{aligned}(J, K) \in \ \text{submap} \quad \Leftrightarrow \quad & (J_T \subseteq K_T) \\ & \wedge (J_R \subseteq K_R) \\ & \wedge (J_A \subseteq K_A) \\ & \wedge (J_C \subseteq K_C) \\ & \wedge (J_Z \subseteq K_Z) \\ & \wedge (J_I \subseteq K_I) \\ & \wedge J, K \in M^*. \end{aligned} \tag{15}$$

Merge is a function from a pair of maps to a merged map.

$$\text{merge} = f : (M \times M) \rightarrow M. \tag{16}$$

Topic maps **merge** by performing the pairwise union of their elements. It is inherent in the definition of sets and elements that assertions are the same if they are indistinguishable; that is, when their types have the same roles, and when their castings have the same role players for those roles. If two assertions can be distinguished from one another, then they are not the same.

$$\begin{aligned}\text{merge}(J, K) = L \quad \Leftrightarrow \quad & (J_T \cup K_T = L_T) \\ & \wedge (J_R \cup K_R = L_R) \\ & \wedge (J_A \cup K_A = L_A) \\ & \wedge (J_C \cup K_C = L_C) \\ & \wedge (J_Z \cup K_Z = L_Z) \\ & \wedge (J_I \cup K_I = L_I) \\ & \wedge J, K, L \in M^*. \end{aligned} \tag{17}$$

5 Representative Graphs

Topic maps are often drawn as graphs, where the vertices of the graph represent the elements of the topic map and the edges of the graph represent the relations within the types, assertions, and castings. Thus "draw" is a function from topic map M to graph G, provided that the subsequent constraints also hold.

$$\text{draw} \; = f : M \to G. \tag{18}$$

Graph G is an ordered pair of vertices and edges.

$$G = (V, E). \tag{19}$$

The set of vertices are the elements of the map.

$$V = M_T \cup M_R \cup M_A \cup M_C \cup M_Z \cup M_I. \tag{20}$$

The set of edges is a binary relation on the vertices.

$$E \subseteq V \times V. \tag{21}$$

Types in T map to edges in E.

$$t = r_1, r_2, ..., r_n \in T \Leftrightarrow (t, r_1), (t, r_2), ..., (t, r_n) \in E. \tag{22}$$

Assertions in A map to edges in E.

$$a = (t, f) \in A \Leftrightarrow (a, t) \in E. \tag{23}$$

Castings in C map to edges in E.

$$\begin{aligned} c = (a, r, x) \in C \Leftrightarrow (a, c), (c, r), (c, x) \in E; \\ c = (a, r) \in C \Leftrightarrow (a, c), (c, r) \in E. \end{aligned} \tag{24}$$

A Change Log

```
[20031014 - NK]
Axiom 3
  Steve requested a textual modification.
  "The set of types is a subset of the all the possible subsets of the set of roles."

[20030916 - NK]
Section 2
  Added sentence explaining how sets can have different types of elements.

Axiom 1 (Definition of M)
  Changed from UNION to SIX-TUPLE
  Given an element of M we can now tell which subset it is in (T,R,A,C,Z,I)
  Elements in the sets of M now have ‘‘types’’

Axiom 6 (Definition of R)
  Made the definition align with the others
  R = \union_{i} t_i, for all t_i \in T

Axiom 8
  In text, changed ‘‘order pair’’ to ‘‘ordered pair’’

Axiom 12
  Inserted the text ‘‘No castings exist otherwise.’’

Axiom 13 (Definition of Z)
  Made the definition align with the others
  Z = { x : ... }
  Updated text to reflect adjusted defintion

Section 4
  Added definition of M* ‘‘the set of all topic maps’’

Equation 15 (Definition of SUBMAP)
  Changed text and defintion of SUBMAP to align with Axiom 1

Equation 17 (Definition of MERGE)
  Changed text and defintion of MERGE to align with Axiom 1

Equation 19 (Definition of G)
  Now G = V,E
  Edges and vertices should be distinct
  One should always be able to tell the edges and vertices apart

Equation 20
  Changed to align with Axiom 1

Equation 22 (Restriction on edges in E)
  Removed this restriction entirely
  In one odd, self-referential case, (c,r) and (r,c) are simultaneously allowed in the graph

Equations 23-25 (Edges in E)
  Changed ‘‘implies’’ to ‘‘if and only if’’
```